



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/087,651

03/01/2002

Brent Townshend

01-185-A

7884

7590 11/27/2007  
Matthew J. Sampson  
McDonnell Boehnen Hulbert & Berghoff  
32nd Floor  
300 S. Wacker Drive  
Chicago, IL 60606

EXAMINER

OPSASNICK, MICHAEL N

ART UNIT

PAPER NUMBER

2626

MAIL DATE

DELIVERY MODE

11/27/2007

PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents  
United States Patent and Trademark Office  
P.O. Box 1450  
Alexandria, VA 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

**MAILED**  
**NOV 27 2007**  
**Technology Center 2600**

**BEFORE THE BOARD OF PATENT APPEALS  
AND INTERFERENCES**

Application Number: 10/087,651  
Filing Date: March 01, 2002  
Appellant(s): TOWNSHEND ET AL.

---

Lisa M. Shoedel  
For Appellant

**EXAMINER'S ANSWER**

This is in response to the appeal brief filed August 13, 2007, appealing from the Office action mailed March 21<sup>st</sup>, 2007.

**(1) Real Party in Interest**

A statement identifying by name the real party in interest is contained in the brief.

**(2) Related Appeals and Interferences**

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

**(3) Status of Claims**

The statement of the status of claims contained in the brief is correct.

**(4) Status of Amendments After Final**

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

**(5) Summary of Claimed Subject Matter**

The summary of claimed subject matter contained in the brief is correct.

**(6) Grounds of Rejection to be Reviewed on Appeal**

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

**(7) Claims Appendix**

The copy of the appealed claims contained in the Appendix to the brief is correct.

**(8) Evidence Relied Upon**

5,634,086	RTISCHEV et al	5-1997
6,122,614	KAHN et al	09-2000
4,783,803	BAKER et al	11-1985
5,059,127	LEWIS et al	10-1991

**(9) Grounds of Rejection**

The following ground(s) of rejection are applicable to the appealed claims:

***Claim Rejections - 35 USC § 103***

1. Claims 1,2,4,5,6,7, 8, 11-18, 24, 26, 28,29, 32, 33, 36,37,42,44 and 45, are rejected under 35 U.S.C. 103(a) as being unpatentable over the combination of Rtischev et al (5,634,086) in view of Kahn et al (6122614) In further view of Baker et al (4783803).

As per claims 1, 2, 5, 7, 36,37 and 45, Rtischev *et al.* teach:

a means for hearing at least one person who is repeating items (spoken-language instruction apparatus employing speech recognition with user reading words from a written script from an inherent database, Abstract; user and microphone, or user and telephone, Fig. 1, elements 12 and 16, or elements 12 and 14, respectively); and

means for preparing a transcription of what was heard by the means for hearing (as input speech being transcribed – col. 5 lines 5-27)

means for comparing the items with a transcription of what was heard and thus measuring intelligibility from the comparison (speech recognizer using nonlinear HMM speech models, Fig. 3, element 112; preselected script, element 114; score set, element 120; reading errors, col. 3, lines 43 and 47).

A measurement unit operable to determine an intelligibility score of the speaker by comparing the items and a transcription of what the listener hears when the speaker repeats the items (as FSM comparing the resubmitted sentence after determining the previous result wasn't satisfactory (col. 6 lines 54-67)

Rtischev *et al.* also teaches evaluations using data selected from at least speaker responses and items (col. 3, lines 11-17).

As per claim 8, the reading errors (col. 3, lines 43 and 47) of Rtischev *et al.* inherently include at least word substitutions, for an error in reading a word could cause the ASR to interpret it as a different existing word (e.g. a Japanese reader using Rtischev *et al.* apparatus to learn English might pronounce frame as flame, which would cause the ASR to recognize the spoken word as the latter).

As per claims 1, 2, 5, 7, 36,37 and 45, Rtischev *et al.* does not explicitly teach a human listener preparing the transcription without prior knowledge of the text that is being spoken, however, Kahn et al (6122614) teaches a human operator transcribing an audio file (col. 8 lines 19-28) and performing a secondary comparison between the audio file and the initial transcription (col. 8 lines 29-35; fig. 2c). Therefore, it would have been obvious to one of ordinary skill in the art of speech transcription to modify the teachings of Rtischev *et al.* with operator based transcription and double-checking because it would advantageously allow for human interpretation of the audio file without tying time resources of the user themselves (Kahn et al, col. 1 lines 17-26).

The combination of Rtischev et al in view of Kahn now teaches a system wherein the user's input is transcribed (Rtischev et al, as input speech being transcribed – col. 5 lines 5-27), the user's input being compared to a transcription (Rtischev, as FSM comparing the resubmitted

sentence after determining the previous result wasn't satisfactory (col. 6 lines 54-67); wherein the one transcription come from a human operator placed between the user and the system (Kahn et al (6122614) teaches a human operator transcribing an audio file (col. 8 lines 19-28; the operator manually inputs the words that the operator "perceives to hear")).

The combination of Rtischev et al in view of Kahn does not explicitly teach the automatic recognition of the operator's interpretation of what is heard (that is, the claim elements require the operator to speak what is heard; whereas the combination of Rtischev et al in view of Kahn requires the operator to type what is heard). Baker et al (4783803), however, teaches a microphonic input for a user to input speech for recognition (Baker et al (4783803), col. 6 lines 20-26), wherein the output is in word text format and is allowed to be edited by the user (Baker et al (4783803), col. 45 line 54 – col. 46 line 4). Therefore, it would have been obvious to one of ordinary skill in the art of speech recognition at the time the invention was made to modify the system combination of Rtischev et al in view of Kahn with automated speech recognition of the operator's interpretation of the user's input because it would advantageously speed up the recording process of the operator's perceived translation (as opposed to typing the translation manually), (Baker et al (4783803), col. 1 lines 13-20).

As per claims 4,6, 8, 11-13, 15-18, 24, 26, 28,29, 32, 33, 42, and 44, the combination of Rtischev et al in view of Kahn et al (6122614) in further view of Baker et al (4783803) teaches listening to the speaker (Kahn et al (6122614), via the recorded audio file of the user – col. 8 lines 19-28).

As per claim 14, Rtischev et al in view of Kahn et al (6122614) in further view of Baker et al (4783803) does not explicitly teach selecting listeners based on certain background characteristics. However, it would have been obvious for an artisan at the time of invention to select listeners that have extensive background speaking knowledge of the language being learned because they would be best able to determine the intelligibility of someone trying to speak the language.

2. Claims 9, 10, 19, 20-23, 34, 35, 38,40-41, and 43 are rejected under 35 U.S.C. 103(a) as being unpatentable over the combination of Rtischev et al in view of Kahn et al (6122614) in further view of Baker et al (4783803), as applied to claims 1, 2, 5, 7, 8, 36, and 37, above, in view of Charles Lewis et al (U.S. Patent 5,059,127).

As per claims 9, 10, 19, 20-23, 25, 34, 35, 41, and 43, while Rtischev et al in view of Kahn et al (6122614) in further view of Baker et al (4783803) teaches evaluating an error count intelligibility score (Rtischev et al ,reading errors, col. 3, lines 43 and 47), they do not evaluate difficulty of the items and ability of a listener, nor doing this using Item Response Theory. However, Lewis et al. do (col. 1, line 63 through col. 2, lines 1, 16-26, and 42-43, with Figure 1A).

It would have been obvious for an artisan at the time of invention to do this because Lewis et al. teach that IRT Aallows creation of a test in which different individuals receive different questions, yet can be scored on a common scale as well as permits determination in advance of test administration of the level of ability and the accuracy with which ability has been measured (col. 2, lines 31-36).



The rest of the limitations were discussed in connection with the rejection of claims 1, 2, 5, 7, 8, 36, and 37, above.

As per claim 38, Rtischev et al in view of Kahn et al (6122614) in further view of Baker et al (4783803) does not teach a database containing data from previous evaluations. However, Lewis *et al.* teach retaining data from previous testlets and sequentially administering testlets ... until a pass/fail decision can be made (Abstract), thus suggesting retaining results of previous intelligibility evaluations (testlets) for later continued evaluation. It would have been obvious for an artisan at the time of invention to do this, to avoid having to administer all the testlets in a single sitting.

As per claim 40, Rtischev *et al.* suggest use of nonlinear artificial neural net models for speech recognition (see reference to Kim *et al.* under Other Publications on the front page, top of second column).

#### **(10) Response to Argument**

Appellant's main arguments can be found on the bottom of page 8 to the top of page 9, wherein Appellant states "Neither Rtischev, nor Kahn, nor Baker teach or suggest:

- (i) a listener who hears a speaker speaking items and then repeats aloud what is heard;
- (ii) automatically preparing a transcription of what the listener repeats aloud;

(iii) comparing the transcription of what the listener repeated aloud with a text of the items spoken by the speaker;

(iv) measuring intelligibility as recited in claims 1,2,4-24,26,28,29,32-38, and 40-45”.

Examiner disagrees.

Before embarking on the rebuttal of the Appellant’s observations with respect to the relied-upon references, the examiner wishes to make clear what the combination of Rtischev in view of Kahn teaches. As will also be explained below with respect to the references:

Rtischev et al (5,634,086) teaches:

a speaker speaking items (*partial claim element (i)*);  
automatically preparing a transcription (*partial claim element (ii)*);  
comparing the transcription with a text of the items spoken by the speaker (*partial claim element (iii)*);  
measuring intelligibility (*claim element (iv)*);

Kahn et al (6122614) teaches:

a listener who hears a speaker speaking items (*partial claim element (i)*);  
preparing a transcription of the listener (*partial claim element (ii)*);

It is imperative to note, at this juncture, that the user of Ritschev et al corresponds to the “speaker” in the claim elements, and that the operator of Kahn et al corresponds to the “listener”

in the claim elements. The combination of Ritschev et al in view of Kahn et al constructs a parallel process in which the speaker's input is 1) automatically transcribed (Ritschev et al); as well as 2) transcribed by a listener (Kahn et al, the operator). (The Baker et al reference completes the combination by allowing for the listener to repeat aloud what is heard).

Addressing Appellant's Interpretation of the References:

**Rtischev et al:**

As per appellants discussion of the Rtischev reference, examiner generally agrees with the page 9 Appeal Brief assessment of the Rtischev reference. Examiner does note that Rtischev does not disclose a human listener who hear the user. Regarding page 10 of the Appeal Brief, examiner disagrees and argues that Rtischev does provide a transcription of the input speech (col. 5 lines 5-27), a comparison of the items with a transcription of the input and measuring intelligibility (as the speech recognized using nonlinear HMM speech models – Fig. 3, element 112, preselected scripts, element 114; scoring/comparison, element 120, and generating reading errors (col. 3 lines 43-47). In summary, Rtischev teaches a user inputting speech, transcribing the speech, performing comparisons to stored/pretrained models, and generating a score (i.e., reading error) which measures the intelligibility of the comparison. What Rtischev teaches (as noted in the Office Action) is underlined as follows:

(i) a listener who hears a speaker speaking items and then repeats aloud what is heard;

- (ii) automatically preparing a transcription of what the listener repeats aloud;
- (iii) comparing the transcription of what the listener repeated aloud with a text of the items spoken by the speaker;
- (iv) measuring intelligibility

**Kahn et al:**

As per the Kahn reference, examiner agrees with appellant's statement that Kahn teaches a human operator manually transcribing the voice dictation file. Appellant then states that "there is no reason, much less any suggestion, for the human operator to repeat aloud what he hears on the voice dictation file". Examiner argues that no where in the Office Action did the examiner make this allegation – in fact, the Examiner introduced the Baker reference to address this limitation. At this point, it is imperative to analyze what the combination of Ritschev in view of Kahn (and motivation to combine provided in the Office Action rejection) teaches (underlined):

- (i) a listener who hears a speaker speaking items and then repeats aloud what is heard;
- (ii) automatically preparing a transcription of what the listener repeats aloud;
- (iii) comparing the transcription of what the listener repeated aloud with a text of the items spoken by the speaker;
- (iv) measuring intelligibility

The combination of Rtischev et al in view of Kahn now teaches a system wherein the user's input is transcribed (Rtischev et al, as input speech being transcribed – col. 5 lines 5-27), the user's input being compared to a transcription (Rtischev, as FSM comparing the resubmitted sentence after determining the previous result wasn't satisfactory (col. 6 lines 54-67); wherein the one transcription come from a human operator placed between the user and the system (Kahn et al (6122614) teaches a human operator transcribing an audio file (col. 8 lines 19-28; the operator manually inputs the words that the operator "perceives to hear")). Appellants arguments against Kahn not teaching other claim elements is moot because the Kahn reference in not relied upon to teach these claim limitations. The purpose of the introduction of the Kahn reference is to have the added benefit of an operator based transcription to allow for human interpretation of the audio file without tying time resources of the user (Kahn et al, col. 1 lines 17-26). The combination of Ritschev in view of Kahn, does not address the listener/operator repeating aloud.

**Baker et al:**

As per the Baker reference, examiner disagrees with the bulk of appellant's statements on pages 13-15 of the Appeal Brief. Appellants arguments against Baker et al not teaching other claim elements is moot because the Baker et al reference in not relied upon to teach these claim limitations. The purpose of the introduction of the Baker et al reference is to teach the listener (i.e., the operator of Kahn, and not the user of Rtischev)

to speak aloud what the operator himself has heard. Baker et al meets the limitation of “the listener repeating aloud what is heard”, by providing an automated speech recognition system for the operator of Kahn, so that the transcription from the operator in Kahn occurs faster (via speech recognition, and not manually typing the transcription. As such, the combination of Ritschev in view of Kahn et al in view of Baker et al meets the claim limitations:

(i) a listener who hears a speaker speaking items and then repeats aloud what is heard;

(ii) automatically preparing a transcription of what the listener repeats aloud;

(iii) comparing the transcription of what the listener repeated aloud with a text of the items spoken by the speaker;

(iv) measuring intelligibility

Remaining Arguments:

Appellants arguments on page 16 of the Appeal Brief rely upon similar arguments presented previously in the Appeal Brief and as such, Examiner points to the rebuttal presented above in addressing these arguments.

Furthermore, examiner rebuts that:

A) the arguments presented on page 9-15 of the response (as well as previous arguments presented) have dissected the references, but have not treated the references as a combination. In response to applicant's arguments against the references individually, one cannot show nonobviousness by attacking references

individually where the rejections are based on combinations of references. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.*, 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986);

B) In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, the motivation to combine the Rtischev and Kahn references (Rtischev *et al* with operator based transcription and double-checking because it would advantageously allow for human interpretation of the audio file without tying time resources of the user themselves (Kahn *et al*, col. 1 lines 17-26) and the combination of Rtischev in view of Kahn and the Baker reference (modifying the system combination of Rtischev *et al* in view of Kahn with automated speech recognition of the operator's interpretation of the user's input because it would advantageously speed up the recording process of the operator's perceived translation - as opposed to typing the translation manually - Baker *et al* (4783803), col. 1 lines 13-20).

**(11) Related Proceeding(s) Appendix**

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.


For the above reasons, it is believed that the rejections should be sustained.


Respectfully submitted,



Michael N Opsasnick  
Primary Examiner AU2626

Conferees:

Michael N Opsasnick, Primary Examiner, AU2626 

Richemond Dorvil, SPE, AU2626 

Patrick Edouard, SPE, AU2626 